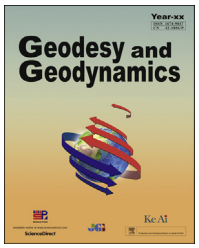
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/en/journals/geog;
http://www.jgg09.com/jweb_ddcl_en/EN/volumn/home.shtml



A regional GNSS-VTEC model over Nigeria using neural networks: A novel approach

Daniel Okoh[a,*], *Oluwafisayo Owolabi*[b], *Christopher Ekechukwu*[c], *Olanike Folarin*[b], *Gila Arhiwo*[d], *Joseph Agbo*[e], *Segun Bolaji*[b], *Babatunde Rabiu*[a]

[a] *NASRDA Center for Atmospheric Research, Kogi State University, Anyigba Campus, Nigeria*
[b] *Ionospheric & Space Physics Laboratory, Department of Physics, University of Lagos, Nigeria*
[c] *Department of Physics, University of Calabar, Nigeria*
[d] *Department of Physics, Federal University of Agriculture, Makurdi, Nigeria*
[e] *Department of Surveying & Geo-informatics, Nnamdi Azikiwe University, Awka, Nigeria*

ARTICLE INFO

Article history:
Received 21 September 2015
Accepted 7 December 2015
Available online 4 April 2016

Keywords:
Global Navigation Satellite System (GNSS) ionosphere
Total electron content (TEC)
Nigerian permanent GNSS network (NIGNET)
Neural network
International reference ionosphere (IRI)

ABSTRACT

A neural network model of the Global Navigation Satellite System – vertical total electron content (GNSS-VTEC) over Nigeria is developed. A new approach that has been utilized in this work is the consideration of the International Reference Ionosphere's (IRI's) critical plasma frequency (foF2) parameter as an additional neuron for the network's input layer. The work also explores the effects of using various other input layer neurons like disturbance storm time (DST) and sunspot number. All available GNSS data from the Nigerian Permanent GNSS Network (NIGNET) were used, and these cover the period from 2011 to 2015, for 14 stations. Asides increasing the learning accuracy of the networks, the inclusion of the IRI's foF2 parameter as an input neuron is ideal for making the networks to learn long-term solar cycle variations. This is important especially for regions, like in this work, where the GNSS data is available for less than the period of a solar cycle. The neural network model developed in this work has been tested for time-varying and spatial performances. The latest 10% of the GNSS observations from each of the stations were used to test the forecasting ability of the networks, while data from 2 of the stations were entirely used for spatial performance testing. The results show that root-mean-squared-errors were generally less than 8.5 TEC units for all modes of testing performed using the optimal network. When compared to other models, the model developed in this work was observed to reduce the prediction errors to about half those of the NeQuick and the IRI model.

* *Corresponding author*. Tel.: +234 8136094616.
E-mail address: okodan2003@gmail.com (D. Okoh).
Peer review under responsibility of Institute of Seismology, China Earthquake Administration.

Production and Hosting by Elsevier on behalf of KeAi

http://dx.doi.org/10.1016/j.geog.2016.03.003

1. Introduction

The ionosphere plays very important roles on our communication systems; when radio waves propagate through the ionosphere, they experience frequency-dependent group delays which become sources of error for the Global Navigation Satellite System (GNSS). Ionospheric delay is particularly a problem for single frequency receivers, which cannot eliminate first-order ionospheric delays by combining observations at two frequencies. Single frequency users rely on applying corrections based on prediction models or on regional models formed based on actual data collected by a network of receivers [1]. Ionospheric delays are proportional to the total electron content (TEC) [2], which represents the number of free electrons contained in a 1 m squared column, along the path of the signal through the ionosphere. This proportionality forms the underlying principle for deriving TEC values from dual-frequency GNSS receivers.

Among other applications, ionospheric TEC models are useful for GNSS error corrections, radio wave propagations, surveying and geo-informatics. Recent explosion of TEC data from the GNSS is spurring interest in using computer neural networks for TEC modeling. Neural networks with their capability for machine learning as well as pattern recognition have been demonstrated to be powerful tools for predictive modeling. A number of works have shown that neural networks (NNs) are good candidates for ionospheric modeling [1,3–6]. NNs operate in a manner that is similar to the human brain; the networks are composed of simple elements operating in parallel and inspired by the biological nervous system. Neural networks can learn trends and patterns in particular data they are given and consequently be able to correctly predict future trends and patterns for the data. A neural network can be trained to perform a particular function by adjusting the value of connections (also called weights) between elements [7]. The true power and advantages of neural networks lies in the ability to represent both linear and non linear relationships directly from the data being modeled. Traditional linear models are simply inadequate when it comes for true modeling data that contains non linear characteristics [8].

In this work, we explore various approaches to TEC modeling using computer neural networks, and present a new method that utilizes the International Reference Ionosphere's (IRI's) critical plasma frequency (foF2) parameter as an additional neuron for the network's input layer.

2. Data and methods

GNSS-VTEC (vertical total electron content) data used in this work are obtained from the Nigerian Permanent GNSS Network (NIGNET, www.nignet.net). Available data from all 14 stations on the network were used in this work. Data from the network is available for years from 2011 to 2015. The stations are geographically located as shown in Fig. 1, and as described in Table 1. Data from 2 stations (FUTA and GEMB) were used for testing purposes, while data from the other 12 stations were used in the training. FUTA and GEMB were selected

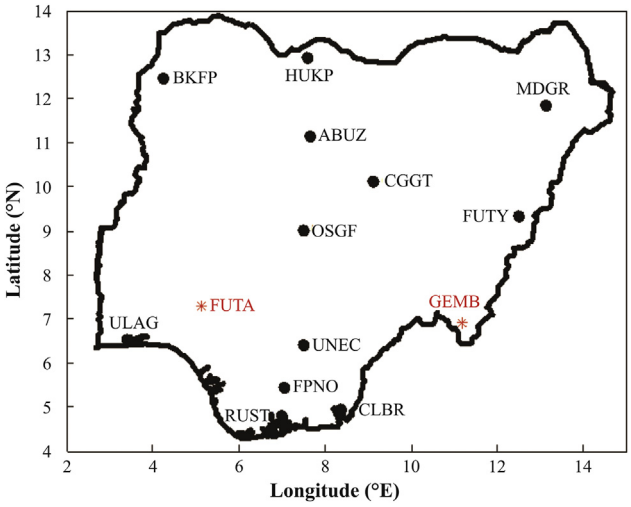


Fig. 1 – Geographical locations of stations on the NIGNET. Stations in red color are used for testing while stations in black color are used in the actual training.

such that they have fewer number of data points, and one of them (FUTA) is inside the area bounded by the stations used in the training while the other (GEMB) is outside of the area.

GNSS data from the NIGNET were obtained in the Receiver Independent Exchange (RINEX) format and processed into VTEC readable formats using software developed by GOPI Krishna Samela. The basic principle harnessed in deriving ionospheric TEC values from GNSS observations is that GNSS signals having different frequencies experience different ionospheric time delays when they transverse the same portion of the ionosphere. Precisely, a GNSS signal of frequency, f, will experience an ionospheric time delay, t, in an amount given by equation (1) [9].

$$t = 40.3 \times \frac{TEC}{cf^2} \tag{1}$$

where c is the speed of light in free space. At two frequencies, f_1 and f_2, the time delay between arrivals of the two signals transmitted the same time from the same GNSS satellite is therefore given by equation (2).

$$\Delta t = 40.3 \times \frac{TEC}{c\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right)} \tag{2}$$

TECs computed using the pseudo-range measurements alone are usually noisy; differential carrier phase measurements are used to obtain precise measures of the relative TECs, and a combination with the pseudo-range measurements provide the absolute slant TEC values (STECs) along the receiver-satellite path ([10–12]). VTECs are obtained from the STECs using equation (3).

$$VTEC = \frac{STEC - (b_r + b_s)}{S(E)} \tag{3}$$

where b_r and b_s are respectively the receiver and satellite biases, $S(E)$ is the mapping function defined by equation (4) [13].

Table 1 – Description of NIGNET stations used in this work (*the sample size is the number of data points after the samples have been averaged in 30-min interval).

Station code	City	State	Latitude (°N)	Longitude (°E)	Elevation (m)	Sample size*
ABUZ	Zaria	Kaduna	11.1517	7.6486	706.1	59,868
BKFP	Birnin-Kebbi	Kebbi	12.4684	4.2292	251.0	67,700
CGGT	Toro	Bauchi	10.1231	9.1181	917.4	20,581
CLBR	Calabar	Cross-River	4.9503	8.3514	61.5	43,156
FPNO	Owerri	Imo	5.4345	7.0331	92.6	7371
FUTA	Akure	Ondo	7.2986	5.1364	416.0	7217
FUTY	Yola	Adamawa	9.3497	12.4978	248.4	63,637
GEMB	Gembu	Taraba	6.917	11.1839	1796.6	5088
HUKP	Katsina	Katsina	12.9211	7.5909	565.0	23,997
MDGR	Maiduguri	Borno	11.8381	13.1309	351.8	16,932
OSGF	Abuja	Federal-district	9.0275	7.4861	533.6	50,638
RUST	Port-Harcourt	Rivers	4.8017	6.9784	46.6	12,804
ULAG	Lagos	Lagos	6.5172	3.3975	45.5	44,841
UNEC	Enugu	Enugu	6.4247	7.5047	255.4	62,175

$$S(E) = \frac{1}{\cos(z)} = \left\{1 - \left(\frac{R_E \times \cos(E)}{R_E + h_S}\right)^2\right\}^{-0.5} \quad (4)$$

where z and E are respectively the zenith and elevation angles in degrees; R_E and h_S are respectively the mean Earth radius and the ionosphere (effective) height above the Earth surface in km. The value of h_S used in this work is 350 km. Other details of the methods and procedures used in developing the GOPI software are contained in reference [10].

To obtain instantaneous values of VTEC for a given location, VTEC values from the various satellites that are visible over the location at the time were averaged excluding those from satellites with elevation angles less than 25°. This is to minimize multipath errors. The resulting VTEC data were further averaged in 30-min intervals to reduce data and to lessen spikes on the data profiles. After averaging, a total of 486005 data points were available as illustrated on the last column of Table 1.

The Levenberg–Marquardt backpropagation algorithm [14] was implemented in this work; weights and bias values were updated according to the optimization method for this algorithm. The algorithm is admired for its speed and efficiency in learning ([15,16]). NNs typically have input layers, output layers, and intermediary hidden layers. Each layer could consist of one or more units (also called neurons).

In this work, we implemented 4 different NN systems with varying input layer neurons. The total number of input layer neurons used is 8 and are as follows: Year, Day of year, Hour of day, Geomagnetic latitude, Geomagnetic longitude, disturbance storm time (DST), sunspot number (SSN), and IRI-foF2. The first 3 neurons are to facilitate the networks' learning of temporal variations; the Year specifically aids learning of solar cycle variations, the Day of year aids learning of seasonal variations, and the Hour of day aids learning of diurnal variations. The fourth and fifth of the neurons are to facilitate the networks' learning of spatial variations. The DST aids their learning of variations associated with geomagnetic activities, while the SSN further supports their learning of solar cycle variations.

The inclusion of IRI-foF2 as an additional input layer neuron is novel in this work. The idea is to increase learning capacity of the networks by using a reliable model. The IRI is an empirical one developed using available data from all around the world, and has been widely accepted as a dependable ionospheric model ([17,18]). The idea of modeled IRI values rather than observed values is to facilitate predictions at all desired locations and at all desired time periods. Further than these considerations, it is foreseen that using a model (like the IRI) which has incorporated long-term solar cycle variation effects will promote the networks' capacity to learn long-term variations. This is specifically useful for regions where data is available for a relatively short period of time. Since the sunspot cycle is about 11 years, approximately an 11-year dataset is required for the network to adequately learn input-target connections that are as results of the solar cycle variations. For regions like Nigeria which have much less than 11 year of observed data, the inclusion of the IRI model as an input layer neuron will further enhance the networks' learning of solar cycle variations.

The implementation of 4 systems of neural network is to evaluate the effects of certain additional input layer neurons on the networks. The difference between the systems is in the set of input layer neurons used, and are as follows:

Network 1: Year, Day of year, Hour of day, Geomagnetic latitude, Geomagnetic longitude.

Network 2: Year, Day of year, Hour of day, Geomagnetic latitude, Geomagnetic longitude, DST, SSN.

Network 3: Year, Day of year, Hour of day, Geomagnetic latitude, Geomagnetic longitude, IRI-foF2.

Network 4: Year, Day of year, Hour of day, Geomagnetic latitude, Geomagnetic longitude, DST, SSN, IRI-foF2.

DST indices were obtained from the World Data Center (WDC) for Geomagnetism (http://wdc.kugi.kyoto-u.ac.jp/dstdir/index.html), while data on SSN were obtained from the WDC-SILSO (Sunspot Index and Long-term Solar Observations, http://www.sidc.be/silso/datafiles), Royal Observatory of Belgium, Brussels.

IRI-foF2 is obtained from the IRI-2012 model. The model is implemented in this work as an executable obtained from the IRI FORTRAN source code; this is to allow for user-defined automatic operations which are also necessary to improve the user-friendliness of the final application developed in this work. The model is used in its standard option mode.

The output layer for each of the networks has one neuron which is the GNSS-VTEC to be modeled. Deciding the number of neurons in the hidden layer is an intricate aspect of neural network trainings, and it is a factor that also affects the performance of the trained networks. The most credible practice has been to train several networks that vary in the number of hidden layer neurons, and then selecting the best of them using a performance index.

In this work, we simulated 30 neural networks for each of the 4 systems of network, varying the number of hidden layer neurons in integer steps from 1 to 30. The main performance index used is the root-mean-squared-errors (RMSEs). Correlation coefficients were also used to ensure that the neural network predictions followed very similar trends as the observations. Our criteria for deciding the best network is to choose the one that consistently gives one of the most minimized RMSEs on various test datasets. Testing of the networks in this work was done on four scenarios using the following test datasets:

1. The latest 10% of the observations from each of the 12 stations used in the actual training.
2. A random 15% of the data remaining after removing the dataset in 1 above. Another random 15% was used for validation.
3. The entire observations from the FUTA station.
4. The entire observations from the GEMB station.

The first two test datasets were used to test the networks' temporal performances while the last two were used to test their spatial performances. The first set was used to test the forecasting performance of the networks, and the second set was used for a general testing of the networks at randomly chosen times. The third set was used to test the networks' performances at locations within the region of stations used in the training, while the fourth set was used to be representative of locations outside of the region. Fig. 2 and 3 (in the next section) show results of the tests, while Fig. 4 shows a schematic of the optimal network used in this work.

3. Results and discussions

Figs. 2 and 3 show correlation coefficient and RMSE profiles for the four different systems of networks implemented in this work, and in each case the number of hidden layer neurons is varied from 1 to 30. The test data used for Fig. 2a and c are the random datasets, while for Fig. 2b and d are the latest datasets. The figures clearly show that network 4 presents the best

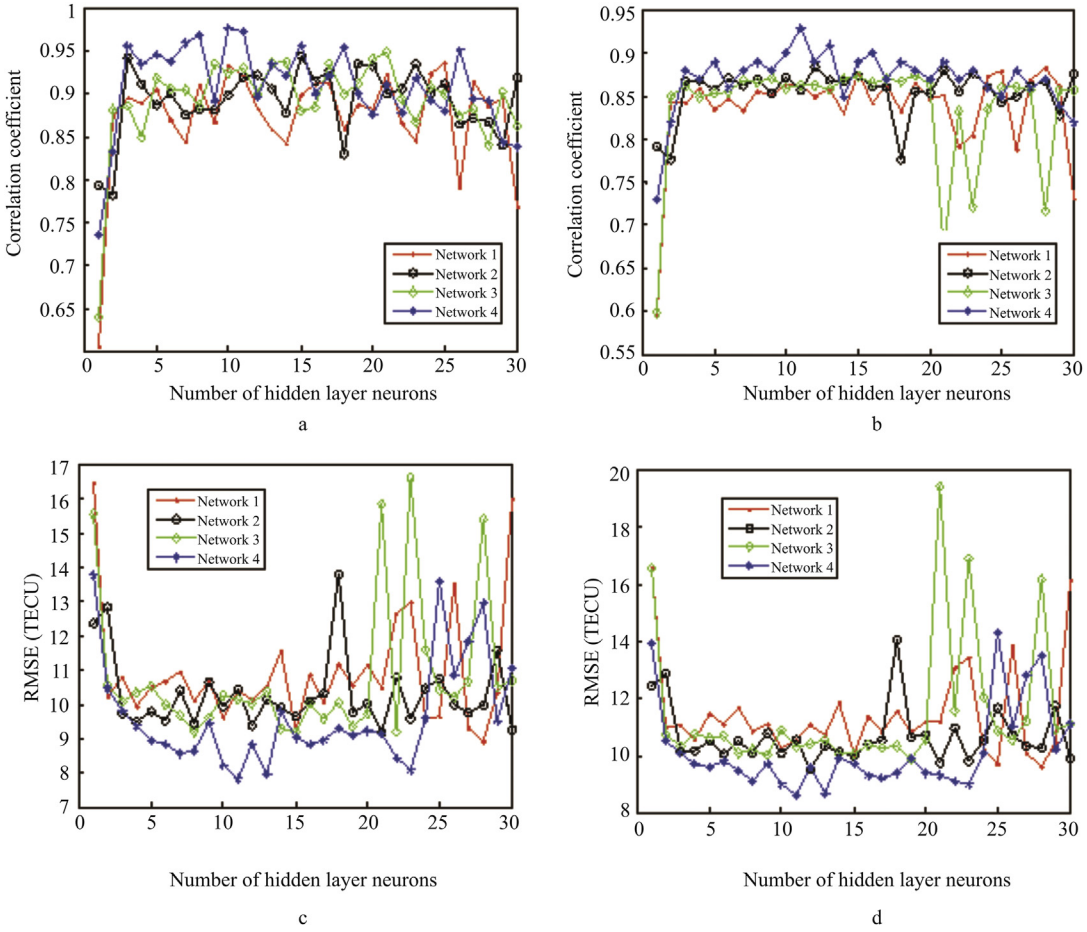


Fig. 2 – Plots of (a) correlation coefficients using the latest dataset, (b) correlation coefficients using the random dataset, (c) RMSE using the latest dataset, (d) RMSE using the random dataset.

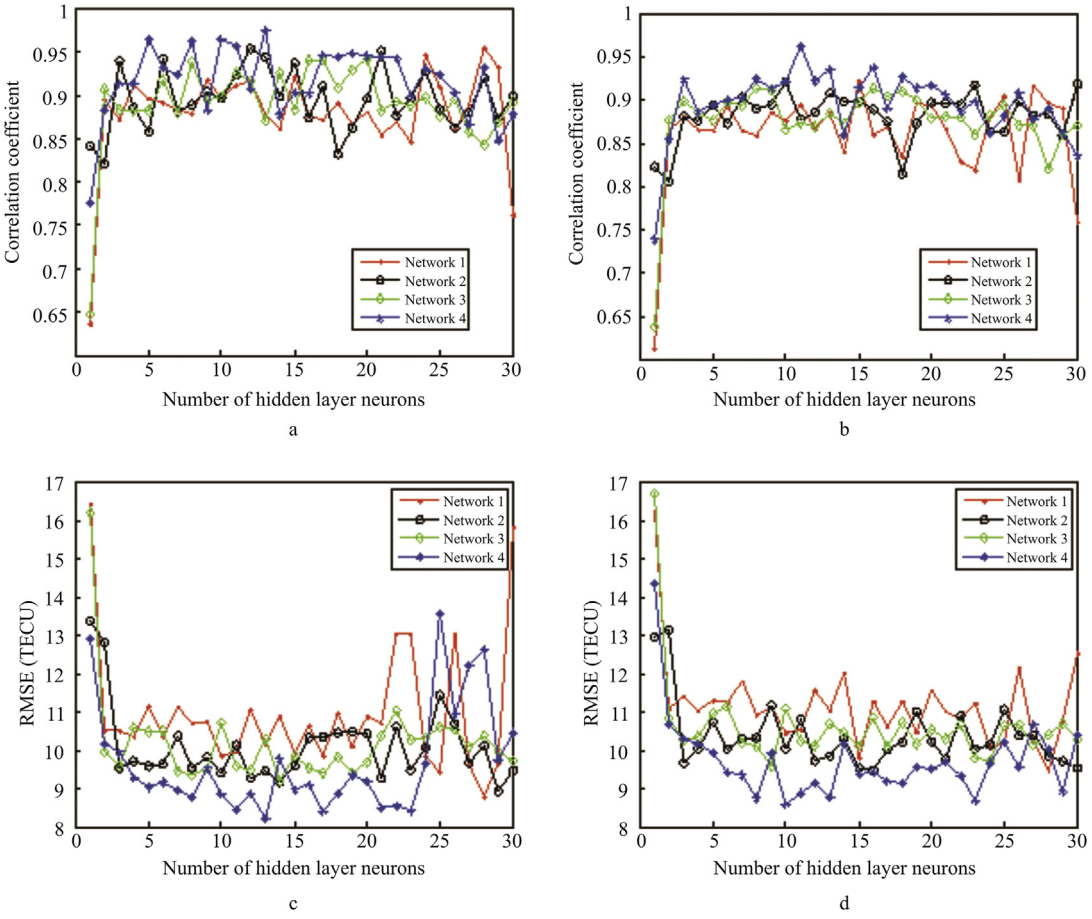


Fig. 3 – Plots of (a) correlation coefficients using the FUTA station, (b) correlation coefficients using the GEMB station, (c) RMSE using the FUTA station, (d) RMSE using the GEMB station.

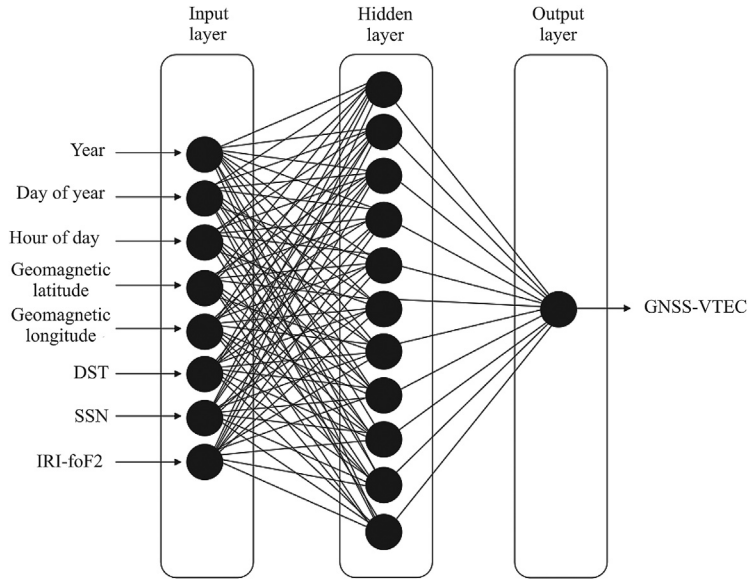


Fig. 4 – Schematic of the optimal neural network implemented in this work.

scenarios. The correlation coefficient plots for network 4 are consistently higher than for the other 3 networks, indicating better agreement in trend between the predictions and the observations when the IRI-foF2 is included as an input layer neuron. The RMSE plots for network 4 are also consistently lower than for the other 3 networks, indicating smaller errors (and so better agreements) between the observations and the network predictions. Network 1 presented the worst case scenarios with lower correlation coefficients and higher RMSEs. This is somewhat expected as it is known that using more functional neurons in the input layer increases the learning capacity of neural networks [19]. It also establishes that the input layer neurons which are additionally included to networks 2, 3 and 4 are really functional as they helped the networks learn better.

Also interesting are the results of the interplay between networks 2 and 3. The major difference is that network 3 uses the IRI-foF2 in place of the DST and SSN used in network 2. While the two networks present relatively close results, it can still be seen that network 3 offers better results especially when the number of hidden layer neurons is less than 20. This is an indication that the IRI-foF2 is such more efficient than the DST and SSN put together. An explanation that justifies this observation is that the IRI-foF2 contains signatures of the DST and SSN amongst other indicators that were incorporated during the development of the IRI model. The standard IRI options used in this work imply that the storm option is turned 'on', this ensures that the model builds in the effects of geomagnetic storm activities to the predicted IRI-foF2 values.

Fig. 3 shows the performance of the networks spatially. Fig. 3a and c use data from the FUTA station while Fig. 3b and d use data from the GEMB station. The FUTA station is within the region of stations used for the training while the GEMB station is outside of the region. Results in Fig. 3 show that the networks' predictions are good for both stations, implying that the networks interpolate and extrapolate well spatially (RMSEs for the optimal networks are generally less 8.5 TECU and the correlation coefficients are generally higher than 0.9). The predictions for the FUTA station are slightly better than for the GEMB station, meaning that the networks interpolate slightly better than they extrapolate. This is also expected since interpolation is done within the limits of data from the stations used in the training, while extrapolation is done on a one-sided reference to the same data.

From the results in Figs. 2 and 3, we establish that network 4 presents the best of the four systems, and that the optimal number of hidden layer neurons is 11. RMSEs and correlation coefficients of this network consistently rank among the best for all the test patterns. This is the network we have adopted for utilization in the final model application developed in this work. We present sample simulations in Figs. 5–7 using the model developed in this work.

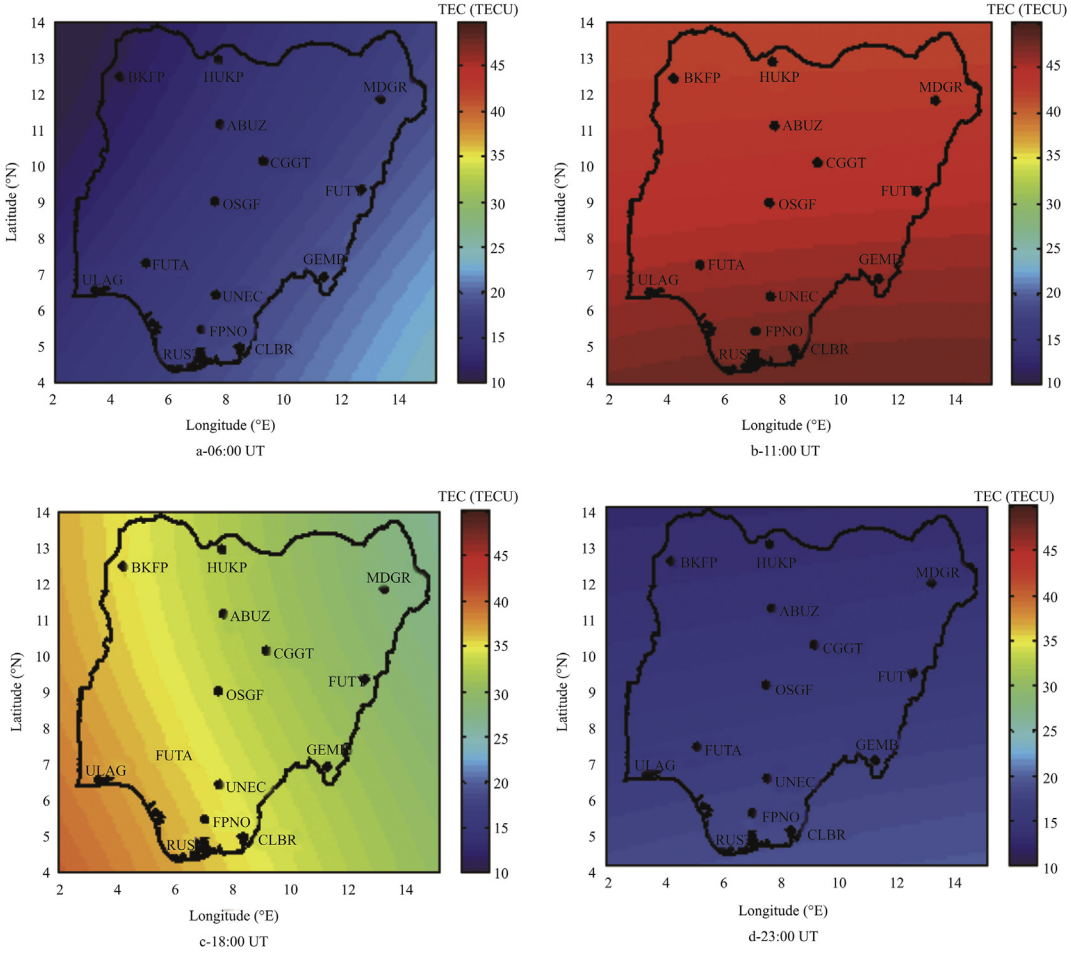


Fig. 5 – VTEC maps for the 183rd day of 2015.

Fig. 5a–d represents the VTEC maps for 06:00 UT (a Nigerian morning), 11:00 UT (the Nigerian midday), 18:00 UT (a Nigerian evening), and 23:00 UT (the Nigerian midnight) respectively of the 183rd day of 2015. The figures indicate very high VTEC values at midday, and reduced values at sunrise and sunset. This clearly supports the already established idea that the Sun is the main source of ionospheric ionization.

Interestingly, the figures also contain the signatures of the Sun's relative westward motion as the Earth rotates eastwards. During the morning hours (Fig. 5a), the Sun can be seen rising from the east as the VTECs are higher eastwards than westwards. The reverse is the case when the Sun sets in the evening hours (Fig. 5c).

Fig. 6a–d represents the VTEC maps for 11:00 UT of the 81st, 171st, 261st, and 351st day respectively of 2015. The days were arbitrarily chosen to represent days in the four seasons of the year (that is, the March equinox, the June solstice, the September equinox, and the December solstice, respectively). Peak ionizations are observed in the March equinox and in the December solstice. During the equinoxes, the Sun is directly overhead the equatorial region than during the solstices. The result is that there is higher photo-ionization (corresponding to higher TEC values) equatorially during the equinoxes than during the solstices. This explains the observed enhanced VTEC during the equinoxes (Fig. 6a and c). The March equinox values are higher than the September equinox values due to the tilt of the Earth's rotational axis on its orbital plane around the Sun. The region of this study is also not centrally located on the geomagnetic equator. This is also going to result in a situation where one of the solstices exhibit higher ionospheric ionization than the other. One of the solstices is usually much hotter than the other (especially and more conspicuous for mid and high latitude regions, but also noticeable in the lower latitudes). The region of this study is geomagnetically located southwards, and the December solstice is the southern summer. This

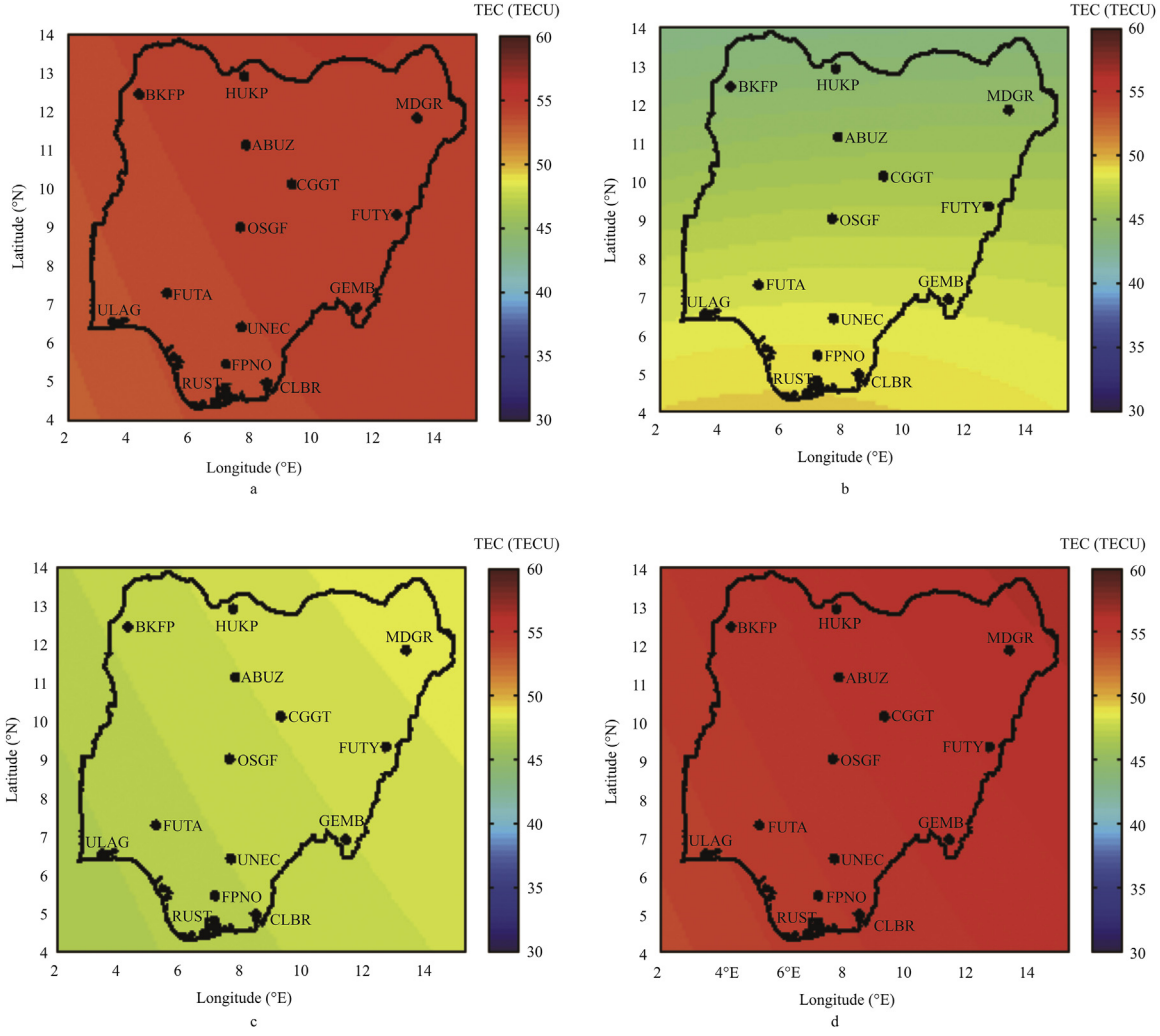


Fig. 6 – VTEC maps for 11:00 UT of (a) the 81st day, (b) the 171st day, (c) the 261st day, (d) the 351st day of 2015.

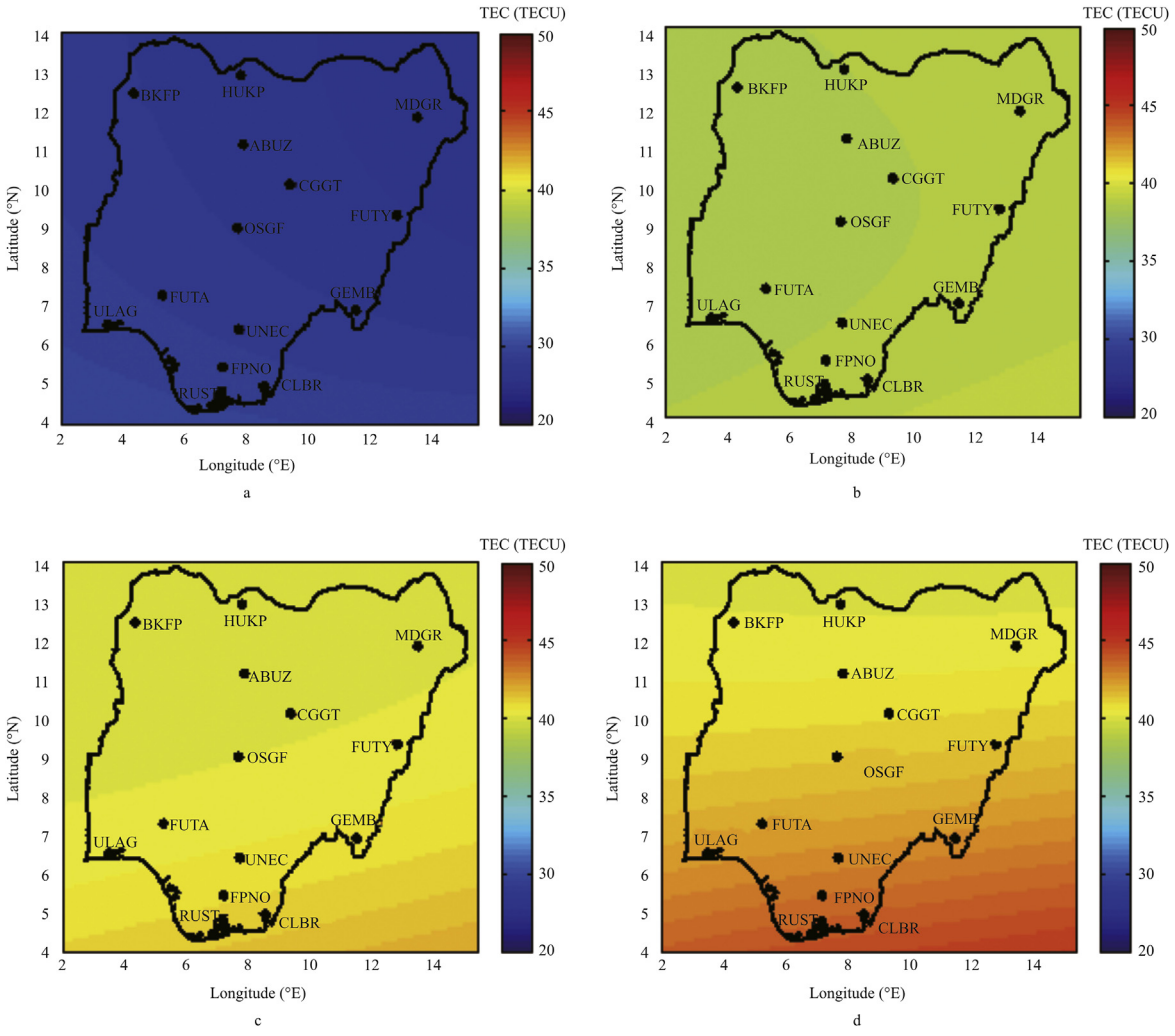


Fig. 7 – VTEC maps for 11:00 UT of the 183rd day of (a) 2011, (b) 2012, (c) 2013, (d) 2014.

explains the observed enhanced VTEC during the December solstice. Similar observations were noted in references [6] and [20].

Fig. 7a–d represents VTEC maps for 11:00 UT of the 183rd day of 2011, 2012, 2013, and 2014 respectively. The scenario is created to observe the models response to long-term variations introduced by the solar activity. The solar activity level has been on the increase between 2011 and 2014. Fig. 7a–d replicates this increase, an indication that the model correctly demonstrates the effects of the long-term solar activity variations on VTEC.

4. Comparison with the IRI model and the NeQuick

To evaluate the performance of the model developed in this work, we present results showing how diurnal VTECs obtained from the model comparing with those of the GNSS observations, alongside those from the IRI model and from the NeQuick.

Three of the NIGNET stations (HUKP, OSGF, and CLBR) were chosen to respectively represent stations in the northern, central, and southern parts of the country. For each of the stations, the comparisons were done using available datasets from 2011 to 2014. In each of the years, four days were chosen such that they are each the first days of the seasons with available GNSS observations. The seasons are described such that days in the months of February, March and April belong to the March equinox, those of May, June and July belong to the June solstice, those of August, September and October belong to the September equinox, and those of November, December and January belong to the December solstice.

RMSEs and correlation coefficients between of the model predictions and the GNSS observations were computed, and the results are as shown in Table 2. In Table 2, the set of results belonging to each day is put in 3 rows; the first row shows the

day, the second row shows the RMSEs, and the third row shows the correlation coefficients. Each of the second and third rows have 3 values separated by commas; the first value is for the model developed in this work, the second value is for the NeQuick, and the third value is for the IRI model. X means there was no GNSS observation for the period. Figs. 8–10 present graphical illustrations of how the model predictions compare with the GNSS observations. In order not to be too cumbersome, only the result of 1 year is presented graphically for each of the stations. Fig. 8 shows the results for the HUKP station in 2014, Fig. 9 shows the results for the OSGF station in year 2013, and Fig. 10 shows the results of the CLBR station in year 2012.

Results from Table 2 and from Figs. 8–10 clearly show very impressive improvements of our model predictions over those of the NeQuick and the IRI model. Statistics of the errors (demonstrated in Fig. 11) show that the VTEC RMSEs in our model is typically about 5 TECU, while for the NeQuick and the IRI model, the values are respectively about 10 TECU and 9 TECU. The worst case RMSE recorded in this comparison in our model is about 10 TECU, as against the 18 TECU and 15 TECU respectively recorded for the NeQuick and the IRI model. The IRI model predictions also seem to be slightly better than those of the NeQuick. It is important to clearly explain that in deriving VTEC from the IRI model, the TEC integration limit was set to 1500 km. This is the farthest that the IRI model is guaranteed to perform well [17]. Ideally, the value should be up to the GPS satellite altitude value of about 20,200 km to make a fitting comparison. This shortfall could account for the observed underestimations in the IRI model predictions. The GPS satellite altitude value of about 20,200 km was used on the NeQuick. Daily f10.7 values from

Table 2 – Statistics of RMSE and correlation coefficient results (X means that there was no GNSS data for the period. For each set of observation, the first row is the day, the second and third rows contain the RMSE and correlation coefficients respectively. Each set of RMSE and correlation coefficient contain 3 values that are separated by commas. The first, second and third RMSE values are respectively the RMSEs between the GNSS-VTEC and the VTECs from this work, from the NeQuick, and from the IRI model. The correlation coefficients are also similarly arranged in that order).

Station	Year	March equinox	June solstice	September equinox	December solstice
HUKP	2011	X	X	X	X
		X	X	X	X
		X	X	X	X
	2012	APRIL 3	MAY 9	X	X
		5.50, 11.55, 8.50	5.52, 8.37, 6.07	X	X
		0.95, 0.93, 0.92	0.99, 0.91, 0.96	X	X
	2013	X	MAY 8	AUGUST 1	NOVEMBER 1
		X	3.94, 10.30, 6.50	6.03, 5.50, 8.00	9.32, 15.28, 13.41
		X	0.98, 0.91, 0.95	0.99, 0.96, 0.98	0.92, 0.79, 0.88
	2014	FEBRUARY 1	MAY 1	AUGUST 31	NOVEMBER 1
		4.68, 10.35, 6.29	6.84, 11.60, 12.43	6.19, 9.48, 11.17	9.20, 16.39, 11.82
		0.97, 0.93, 0.97	0.96, 0.85, 0.89	0.98, 0.94, 0.91	0.99, 0.92, 0.96
OSGF	2011	MARCH 9	MAY 1	AUGUST 1	NOVEMBER 1
		3.53, 7.71, 9.70	3.16, 10.63, 8.95	3.71, 4.51, 3.14	6.81, 13.64, 15.55
		0.97, 0.94, 0.91	0.99, 0.93, 0.96	0.97, 0.96, 0.95	0.96, 0.89, 0.92
	2012	FEBRUARY 1	MAY 1	AUGUST 1	NOVEMBER 1
		4.07, 7.20, 5.63	4.39, 7.71, 6.69	5.57, 6.91, 8.31	3.97, 12.61, 7.07
		0.96, 0.92, 0.94	0.99, 0.92, 0.95	0.99, 0.94, 0.96	0.98, 0.94, 0.94
	2013	FEBRUARY 1	MAY 2	AUGUST 1	NOVEMBER 6
		3.09, 6.03, 5.36	5.17, 12.84, 10.07	2.92, 6.32, 5.41	7.41, 12.36, 11.08
		0.98, 0.89, 0.94	0.98, 0.85, 0.93	0.98, 0.95, 0.95	0.96, 0.95, 0.92
	2014	FEBRUARY 1	X	X	X
		4.32, 9.63, 5.43	X	X	X
		0.99, 0.93, 0.98	X	X	X
CLBR	2011	X	X	X	DEC. 4
		X	X	X	6.03, 14.41, 11.15
		X	X	X	0.97, 0.84, 0.89
	2012	FEBRUARY 1	JULY 11	AUGUST 1	NOVEMBER 1
		3.60, 6.30, 5.32	6.46, 14.03, 8.06	4.54, 11.51, 7.56	3.43, 10.23, 7.67
		0.97, 0.93, 0.95	0.97, 0.87, 0.96	0.99, 0.92, 0.96	0.98, 0.93, 0.95
	2013	FEBRUARY 1	MAY 1	OCTOBER 9	NOVEMBER 1
		3.45, 6.54, 8.74	6.15, 18.20, 11.44	5.80, 11.15, 9.72	8.68, 18.88, 15.26
		0.97, 0.88, 0.93	0.97, 0.86, 0.92	0.96, 0.86, 0.91	0.92, 0.77, 0.85
	2014	FEBRUARY 1	JUNE 28	AUGUST 31	NOVEMBER 1
		4.71, 18.03, 7.80	8.23, 3.88, 15.52	3.38, 7.56, 11.74	10.80, 15.37, 14.79
		0.97, 0.84, 0.96	0.98, 0.97, 0.95	0.99, 0.93, 0.92	0.97, 0.87, 0.90

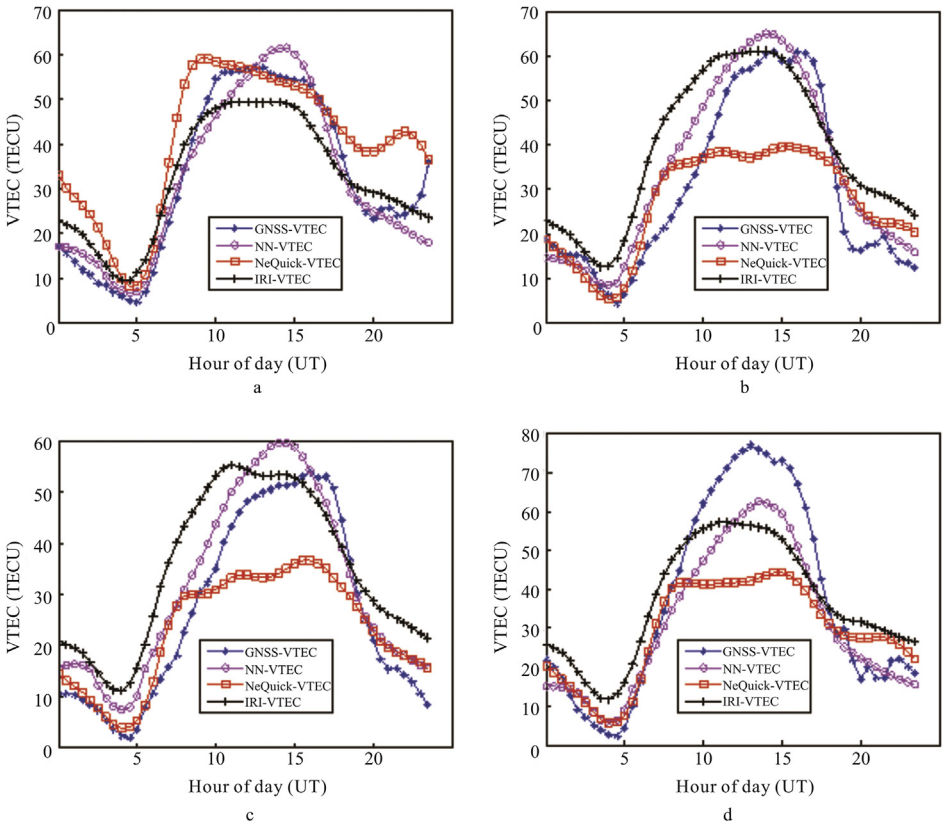


Fig. 8 – Comparison of diurnal VTECs from this work with VTECs from the IRI and NeQuick for days in (a) March equinox, (b) June solstice, (c) September equinox, (d) December solstice, of 2014 for the HUKP Station.

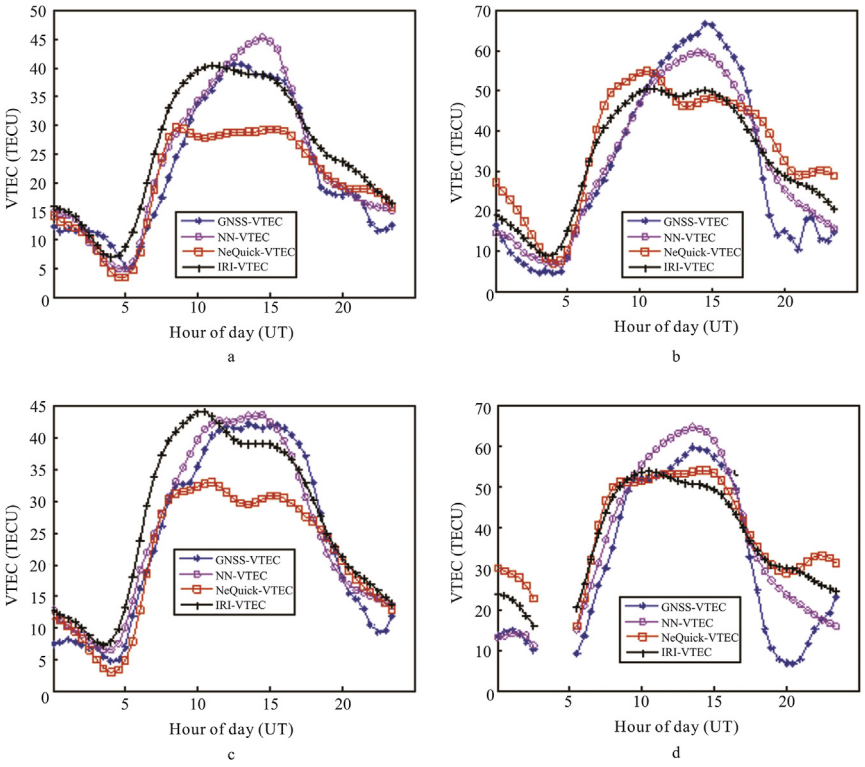


Fig. 9 – Comparison of diurnal VTECs from this work with VTECs from the IRI and NeQuick for days in a) March equinox, b) June solstice, c) September equinox, d) December solstice, of 2013 for the OSGF Station.

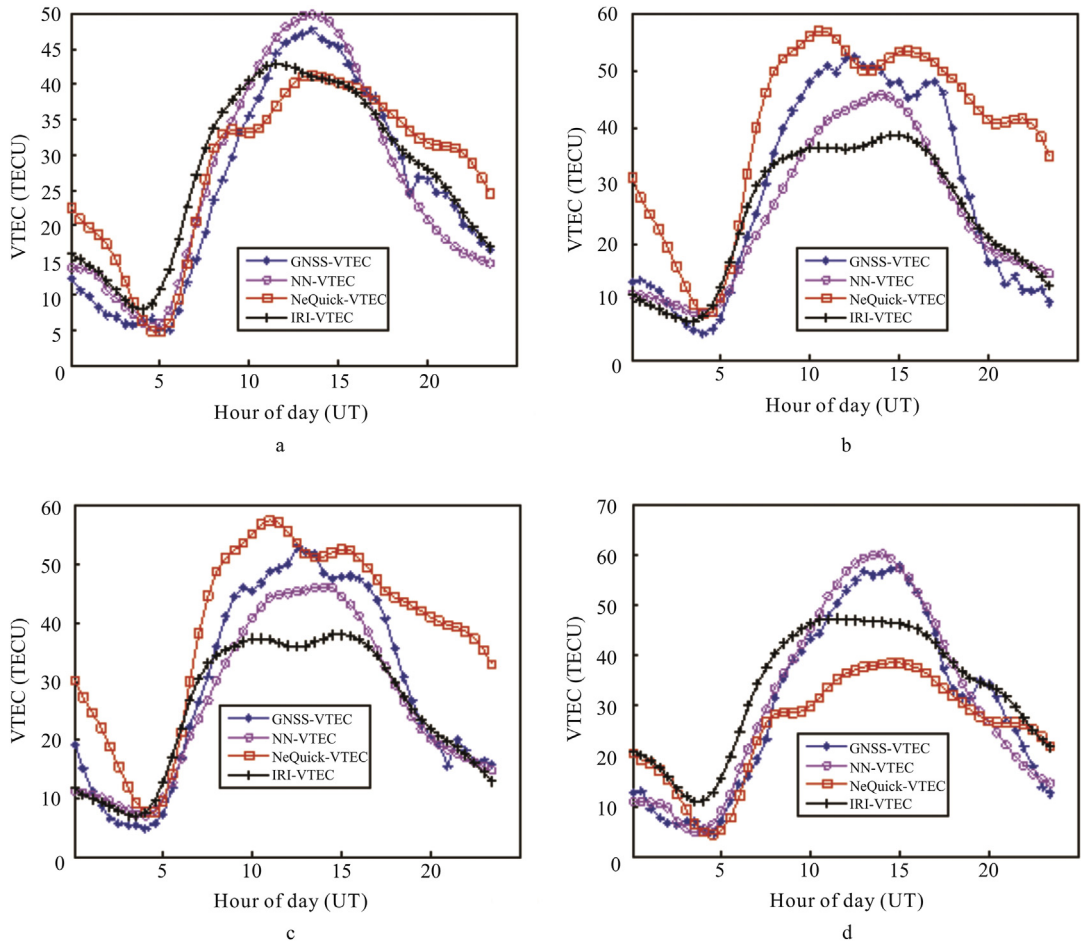


Fig. 10 – Comparison of diurnal VTECs from this work with VTECs from the IRI and NeQuick for days in (a) March equinox, (b) June solstice, (c) September equinox, (d) December solstice, of 2012 for the CLBR station.

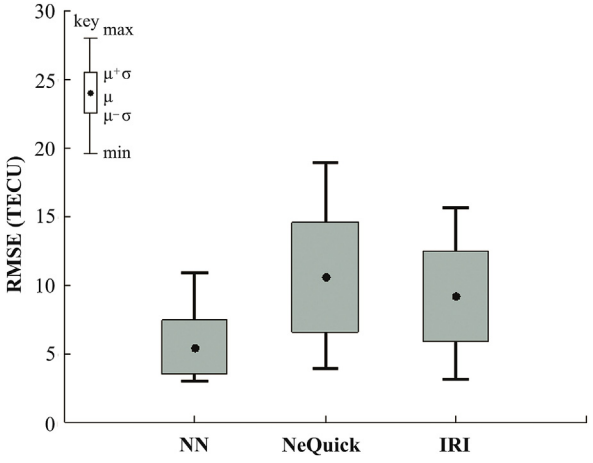


Fig. 11 – Box and Whisker plot showing statistics of the RMSEs.

the NOAA (National Oceanic and Atmospheric Administration, ftp://ftp.ngdc.noaa.gov) were used as solar activity indices on the NeQuick. There is indication from developers of the NeQuick that the f10.7 parameter can be tuned differently for different regions, and this can lead to improvements in the NeQuick predictions for the different regions. Results from this work do not undermine the robustness and the usefulness of global models like the IRI and NeQuick, but rather indicate that these models could serve as reliable backbones for developing better regional models.

5. Conclusion

The paper presents an evaluation of the performance of four neural network systems that vary in the neurons used on the input layer. The evaluation was done to observe the effects of using (or not using) DST, SSN, and IRI-foF2 as input layer neurons on the networks. Results from the work indicate that

all three parameters are effective in increasing the network performances when used as neurons on the input layer. The inclusion of the IRI-foF2 which is a novel concept in this work is the most effective.

Training was done using GNSS VTEC data derived from receivers on the NIGNET. The networks were tested for spatial performance using data from 2 of the stations that were not used for the training. Asides the general testing for temporal performance of the networks, the networks were also tested for forecasting abilities using the latest 10% of datasets obtained from the each station. RMSEs were generally less than 8.5 TECU for all kinds of testing performed on the optimal networks, and correlation coefficient values were higher than 0.9.

The optimal network developed in this work has 8 input layer neurons, 11 hidden layer neurons, and 1 output layer neuron. Sample simulations using the network correctly demonstrate the variations of the VTEC on diurnal, seasonal, and long-term solar cycle scales.

A comparison of VTEC predictions from the model developed in this work with those from the NeQuick and the IRI model show that the predictions from the model in this work are better in terms of the prediction errors; the model developed in this work typically reduces the prediction errors to about half the values obtained in using the NeQuick and IRI model. This does not suggest that the global models are not useful, but rather (as demonstrated in this work) shows that they can play very active roles in developing better regional models.

Acknowledgments

We acknowledge the Office of the Surveyor General of the Nigerian Federation for making GNSS data available from the NIGNET. We appreciate the WDC for Geomagnetism from where we obtained DST data. We also appreciate the WDC-SILSO (Royal Observatory of Belgium, Brussels) from where we obtained the SSN data. Special thanks to the IRI team for making the model available, and to the NOAA from where we obtained f10.7 values. We are immensely grateful to Dr. John Bosco Habarulema for his technical suggestions.

REFERENCES

[1] Leandro RF, Santos MC. A neural network approach for regional vertical total electron content modeling. Studia Geophysica et Geodaetica 2007;51:279–92.
[2] Hofmann-Wellenhof B, Lichteneeger H, Collins J. Global positioning system: theory and practice. Wien, New York: Springer-Verlag; 2001.
[3] Hernandez-Pajares M, Juan J, Sanz J. Neural network modeling of the ionospheric electron content at global scale using GPS. Radio Sci 1997;32:1081–90.
[4] Tulunay E, Senalp ET, Cander LR, Tulunay YK, Bilge AH, Mizrahi E, et al. Development of algorithms and software for forecasting, nowcasting and variability of TEC. Ann Geophys 2004;47:1201–14.
[5] Senalp ET, Tulunay E, Tulunay Y. Total electron content (TEC) forecasting by Cascade modeling, a possible alternative to the IRI-2001. Radio Sci 2008;43:RS4016. http://dx.doi.org/10.1029/2007RS003719. 2008.
[6] Habarulema JB. A contribution to TEC modelling over Southern Africa using GPS data [PhD Thesis]. Rhodes University; 2010.
[7] Mark B. Neural network toolbox for use with MATLAB. USA: The Mathworks Inc; 2002.
[8] Baboo SS, Shereef KI. An efficient weather forecasting system using artificial neural network. Int J Environ Sci Dev 2010;1(4).
[9] Klobuchar JA. Ionospheric effects on GPS. In: Parkinson BW, Spilker JJ, editors. Global positioning system: theory and applications, vol. 2. Progress in Astronautics and Aeronautics; 1996. p. 164.
[10] Rao PVS, Gopi KS, Niranjan K, Prasad SVVD. Temporal and spatial variations in TEC using simultaneous measurements from the Indian network of receivers during the low solar activity period of 2004–2005. Ann Geophys 2006;24:3279–92.
[11] Jin SG, Occhipinti G, Jin R. GNSS ionospheric seismology: recent observation evidences and characteristics. Earth-Sci Rev 2015;147:54–64.
[12] Zhu FY, Wu Y, Lin J, Zhou Y. Temporal and spatial characteristics of VTEC anomalies before Wenchuan Ms8.0 earthquake. Geod Geodyn 2010;1(1):23–8.
[13] Mannucci AJ, Wilson BD, Edwards CD. A new method for monitoring the earth's ionospheric total electron content using the GPS global network, Proc. of ION GPS-93. Inst Navigation 1993:1323–32.
[14] Levenberg K. A method for the solution of certain non-linear problems in least squares. Q Appl Math 1944;2:164–8.
[15] Demuth H, Beale M. Neural network toolbox for use with MATLAB. Mathworks Inc; 2002.
[16] Kisi O, Uncuoglu E. Comparison of three back-propagation training algorithms for two case studies. Indian J Eng Mater Sci 2005;12:434–42.
[17] Bilitza D, McKinnell LA. International reference ionosphere, IRI-2011. In: IRI 2011 Workshop Presentation, SANSA space science, 10–14 October, Hermanus, South Africa; 2011.
[18] Bilitza D, Reinisch BW. International reference ionosphere 2007: improvements and new parameters. Adv Space Res 2008;42(4):599–609.
[19] Ng A. Coursera video lecture on machine learning. 2012. Retrieved 30th August 2012 from, https://class.coursera.org/ml-2012-002/class/index.
[20] Okoh DI, McKinnell LA, Cilliers PJ, Okeke PN. Using GPS-TEC to calibrate VTEC computed with the IRI model over Nigeria. Adv Space Res 2013;52:1791–7.

Dr. Daniel Okoh is a research scientist with the Center for Atmospheric Research, National Space Research and Development Agency, Nigeria. He presently coordinates activities of the Space Environment Research Laboratory.

Daniel holds a First Class BSc Hons Degree in Physics and Astronomy from the University of Nigeria, and a Distinction MSc Degree in Space Physics from Rhodes University. He has also recently received his PhD in Astrophysics from the University of Nigeria. Daniel has received research fellowships and scholarships from prestigious organizations like the World Academy of Science (TWAS), the Chinese Academy of Science (CAS), and the National Astrophysics and Space Science Program (NASSP).

His research interests are in the fields of Space Science and Astrophysics, and he particularly has huge interest in Ionospheric Modeling. He has contributed a number of novel techniques for ionospheric modeling. Daniel enjoys sharing his experiences on how space applications work, and he is delighted to discuss emerging ideas that are innovative for space applications.